Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

ON GAUSS'S METHOD OF SUBSTITUTION.

By Prof. W. Woolsey Johnson, Annapolis, Md.

In Gauss's notation the normal equations arising in the method of Least Squares, for μ unknown quantities $x, y, z, \ldots, t$, are

$$\left.\begin{array}{l} [aa]x + [ab]y + [ac]z + \ldots + [al]t = [an] \\ [ab]x + [bb]y + [bc]z + \ldots + [bl]t = [bn] \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ [al]x + [bl]y + [cl]z + \ldots + [ll]t = [ln] \end{array}\right\}, \tag{1}$$

and the result of eliminating x by means of the first normal equation are written

$$\left.\begin{array}{l} [bb, 1]y + [bc, 1]z + \ldots + [bl, 1]t = [bn, 1] \\ [bc, 1]y + [cc, 1]z + \ldots + [cl, 1]t = [cn, 1] \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \\ [bl, 1]y + [cl, 1]z + \ldots + [ll, 1]t = [ln, 1] \end{array}\right\}. \tag{2}$$

The fact that these "reduced normal equations," like the normal equations (1), have a symmetrical determinant suggests the question, May they not be regarded as normal equations derivable from certain fictitious observation equations for the $\mu - 1$ unknown quantities $y, z, \ldots, t$?

Oppolzer in fact raises the question directly ("Lehrbuch zur Bahnbestimmung der Kometen und Planeten," vol. ii, p. 331) "whether the newly introduced symbols $[bb, 1]$, $[bc, 1]$, ... may be regarded, in analogy with the symbols $[aa]$, $[ab]$, ..., as sums of products of quantities related to each other in the same way;" and he answers this question by finding $\frac{1}{2}m(m-1)$ quantities (where m is the number of observation equations) corresponding to each letter $b, c, \ldots, l$, of which $[bb, 1]$, $[bc, 1]$, ... are the sums of products, (and of squares when the letters of the symbol are alike). From this he proceeds to draw "the important inference that $[bb, 1]$, $[cc, 1]$, etc. are positive.

The fictitious observation equations which most naturally suggest themselves would seem to be the m equations which result from the elimination of x from the observation equations by means of the normal equation for x. That these equations, in fact, fulfil the requirements is readily seen; for,

writing them in the form

$$\left.\begin{array}{l} b'_1 y + c'_1 z + \ldots + l'_1 t = n'_1 \\ b'_2 y + c'_2 z + \ldots + l'_2 t = n'_2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ b'_\mu y + c'_\mu z + \ldots + l'_\mu t = n'_\mu \end{array}\right\}, \tag{3}$$

their coefficients and absolute terms are of the form

$$b' = b - \frac{[ab]}{[aa]} a, \qquad c' = c - \frac{[ac]}{[aa]} a, \qquad \ldots;$$

whence, for example,

$$[b'c'] = \Sigma bc - \frac{[ab]}{[aa]} \Sigma ac - \frac{[ac]}{[aa]} \Sigma ab + \frac{[ab]\,[ac]}{[aa]^2} \Sigma a^2$$

$$= [bc] - \frac{[ab]\,[ac]}{[aa]} = [bc, 1]. \tag{4}$$

So far as I know, this method of viewing the reduced normal equations did not appear explicitly in any treatise upon Least Squares prior to the third edition of W. Jordan's Vermessungskunde,* for reference to which I am indebted to Mr. Harold Jacoby. In the first volume of this work, which bears the separate title "Ausgleichungs–Rechnung," Jordan, at page 64, calls the equations of the form

$$v = ax + by + cz + \ldots + lt - n \tag{5}$$

"Fehlergleichungen," and the results of eliminating x by means of the normal equation for x, in the form

$$v = b'y + c'z + \ldots + l't - n', \tag{6}$$

"Reduzierte Fehlergleichungen." He proceeds to show, as above, that the reduced normal equations are the normal equations which would arise from these reduced error equations, or the corresponding observation equations (3).† Since we can proceed in the same manner to the second reduced error equations and normal equations, and so on, it is obvious that we thus prove directly that

* The first edition appeared under the title "Taschenbuch der praktischen Geometrie," Stuttgart, 1874. The second edition was called "Handbuch der Vermessungskunde," Stuttgart, 1877. The third edition, improved and enlarged, appeared in 1888.

† The idea appears in a less explicit form in Herr und Tinter's "Lehrbuch der Sphärischen Astronomie," 1887, p. 47.

$[vv] = [nn, \mu]$, without the intervention of the equation

$$[vv] = [nn] - [an]x - [bn]y - \ldots - [ln]t\,.$$

It would seem, however, that this aspect of the reduced normal equations must have been present to the mind of Gauss while writing Art. 182 of the "Theoria Motus," in which he discusses the function W which is $[vv]$ considered as a quadratic function of $x, y, z, \ldots, t$, or Σv^2 when the v's are in the form (5).* If $X = 0$ is the normal equation for x, it is shown that $W_1 = W - \frac{X^2}{[aa]}$ is a function independent of x; and, since when $X = 0$ it has the same value as W, it is the result of eliminating x from W by means of the first normal equation. It is then shown that, $Y_1 = 0$, $Z_1 = 0$, etc. being the reduced normal equations, Y_1, Z_1, etc. have the same relation to W_1 that X, Y, etc. have to W. In like manner $W_2 = W_1 - \frac{Y_1^2}{[bb, 1]}$ is independent of y, and so on. Supposing there are four unknown quantities, the function W is thus finally reduced to the form

$$W = \frac{X^2}{[aa]} + \frac{Y_1^2}{[bb, 1]} + \frac{Z_2^2}{[cc, 2]} + \frac{T_3^2}{[dd, 3]} + \text{constant.}\dagger$$

In the "Disquisitio de Elementis ellipticis Palladis" a somewhat different investigation of this function is given in which the characteristic Gaussian notation is used for the first time, and the constant is shown to be $[nn, \mu]$. Gauss here for brevity suppresses the demonstration that $[bb, 1]$, $[cc, 2]$, etc. as well as $[aa]$ are positive, which is necessary to the proof that $[nn, \mu]$ is the minimum value of $[vv]$. But in the "Theoria Motus" he proves, for another purpose, that $[bb, 1]$ is a sum of squares, "for," he says, "it is evident that W_1 is derived from $[vv]$ the quantity x being eliminated from $v_1, v_2, \ldots$ by means of the equation $X = 0$; hence $[bb, 1]$ will be the sum of the coefficients

* Slight changes have been made in the notation.

† This result is then used in finding the probability of a value of t. In the original edition of the "Theoria Motus" it is said that the probability that s (the fourth variable) shall differ from its most probable value by the quantity σ is proportioned to $e^{-\frac{hh\sigma\sigma}{\delta'''}}$ (δ''' here standing for $[dd, 3]$). This was a mistake for $e^{-hh\delta'''\sigma\sigma}$, and the same mistake is repeated two lines below where $\sqrt{\frac{1}{\delta'''}}$ instead of $\sqrt{\delta'''}$ is stated to be the measure of the relative precision to be attributed to the determination. These mistakes are corrected in Bertrand's tranlation, entitled "Methode des Moindres Carrés," p. 129, except that by a misprint the δ has dropped out of the exponential leaving the meaningless expression $e^{-hh'''\sigma^2}$. By a remarkable coincidence we find also in Davis's translation of the "Theoria Motus" $e^{-hh'''\sigma\sigma}$. Indeed there seems to have been a strange fatality connected with this passage; for in the "Gesammelte Werke" the exponential is printed $e^{-hh\delta'''s}$, introducing two new misprints.

of y^2 in $v_1^2, v_2^2, \ldots$ after the elimination, each of which is a square." Again Encke, in his elaboration of Gauss's method (*Berliner Jahrbuch* for 1835, p. 277) remarks, for the same purpose, that "if in each v the value is given to x for which $X = 0$, the sum of the squares of these v's is $W - \frac{X^2}{[aa]}$, and therefore the coefficient of y^2, which is $[bb, 1]$, will be simply a sum of squares."

In other words, W_1 is the value of Σv^2 when each v is expressed in the form (6), so that the coefficient of y^2 in it is $\Sigma b'^2$ or $[b'b']$, which is therefore equivalent to $[bb, 1]$.

Among the advantages of this point of view in treating the reduced normal equations, we may mention not only the ready proof that $[bb, 1]$, $[cc, 2]$, etc. are positive, and that $[nn, \mu]$ is the value of $[vv]$ when $v_1, v_2, \ldots$ are the residuals, but the simplification it introduces into Gauss's elegant discussion of the general value of the sum of the squares of the errors, and also its bearing upon the weights of the unknown quantities.

It is noteworthy that if, after solving the μ normal equations obtained from m observation equations, we should discover that the μ unknown quantities were connected by a rigorous equation of condition which had previously been overlooked, the most probable values of the unknown quantities would be changed *even if the values we had already found for the unknown quantities happened to exactly satisfy the rigorous equation of condition.* At first sight the rigorous relation being exactly satisfied would seem to be a confirmation of the values already obtained; but this is evidently not true, the normal equations used in obtaining them have now no standing, the true process being to eliminate one of the unknown quantities from the observation equations by means of the rigorous condition, and then form $\mu - 1$ normal equations for the remaining unknown quantities. We thus have a system of μ equations for the μ quantities only one of which is by hypothesis satisfied by the values already obtained.

But suppose the rigorous equations of condition happened to coincide with one of the normal equations first found, say that for x. Then, in accordance with what has been shown above, the new normal equation, being formed from observation equations identical with the fictitious observation equations described above (x being the unknown quantity first eliminated) are the same as the reduced normal equations. They therefore lead to the same values of $y, z, \ldots, t$ as before, and hence also to the same value of x.

Moreover we are led also to the same values of the weights of $y, z, \ldots, t$.*

* The value of $[vv]$ remains also unchanged, but the mean error of an observation is decreased, since there are now but $\mu - 1$ unknown quantities and we have to divide $[vv]$ by $m - \mu + 1$ instead of $m - \mu$.

The weight of x is however increased. For supposing, as in the usual process for obtaining the weights, the values of $x, y, \ldots, t$ when fully expanded in terms of the n's to be

$$x = \alpha_1 n_1 + \alpha_2 n_2 + \ldots + \alpha_m n_m ,$$

$$y = \beta_1 n_1 + \beta_2 n_2 + \ldots + \beta_m n_m ,$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

the inverse weight of x is $\Sigma\alpha^2$. The normal equation for x may be written

$$x = \frac{[an]}{[aa]} - \frac{[ab]}{[aa]} y - \frac{[ac]}{[aa]} z - \ldots ;$$

whence equating the coefficients of n_1,

$$\alpha_1 = \frac{a_1}{[aa]} - \frac{[ab]}{[aa]} \beta_1 - \frac{[ac]}{[aa]} \gamma_1 - \ldots .$$

Now if the value of x above is derived from a rigorous equation, the first term is independent of the observed values, and they only enter the value of x through the values of $y, z, \ldots$; whence the coefficient of n_1 becomes

$$\alpha'_1 = - \frac{[ab]}{[aa]} \beta_1 - \frac{[ac]}{[aa]} \gamma_1 - \ldots .$$

and the inverse weight becomes $\Sigma\alpha'^2$. We have, then,

$$\alpha_1 = \frac{a_1}{[aa]} + \alpha'_1 ;$$

whence

$$\Sigma\alpha^2 = \frac{\Sigma a^2}{[aa]^2} + 2 \frac{\Sigma a\alpha'}{[aa]} + \Sigma\alpha'^2$$

$$= \frac{1}{[aa]} + \Sigma\alpha'^2 ,$$

since $\Sigma a\alpha'$ consists of terms involving $\Sigma a\beta$, $\Sigma a\gamma$, etc. each of which is known to vanish. Thus the inverse weight is less and the weight greater than before.